

Interior Penalty Functions and Duality in Linear Programming

I. I. Eremin[†] and L. D. Popov¹

Received February 25, 2012

Abstract—Logarithmic additive terms of barrier type with a penalty parameter are included in the Lagrange function of a linear programming problem. As a result, the problem of searching for saddle points of the modified Lagrangian becomes unconstrained (the saddle point is sought with respect to the whole space of primal and dual variables). Theorems on the asymptotic convergence to the desired solution and analogs of the duality theorems for the arising optimization minimax and maximin problems are formulated.

Keywords: linear programming, duality, inner penalty functions.

DOI: 10.1134/S0081543813090058

INTRODUCTION

Duality issues are central to problems of mathematical programming [1,2]. In the present paper, we attempt to construct asymptotic duality schemes for linear programming problems based on the classical Lagrangian and exterior penalty functions. These functions were included in the computational toolkit of mathematical programming a long time ago (see [3–8] and other papers). They are used to reduce a constrained extremal problem to a (in general, infinite) series of single-type unconstrained minimization problems for a weighted sum of the original target function and an interior penalty function. Under a certain change of the penalty parameter, the solution of the unconstrained problem converges to the required solution of the constrained problem. In the present paper, we consider the Lagrange function of a linear programming problem with the aim of investigating the issues of the asymptotic reduction of problems in which saddle points of convex–concave functions are sought with respect to some constructively described domain (in our case, the nonnegative orthant of an Euclidean space) to a series of parametric problems in which saddle points of some auxiliary functions are sought without such constraints. Auxiliary functions are based on the original function and include additional terms—logarithmic barriers. We formulate convergence theorems for a method that is similar to modern versions of the central path method as well as analogs of duality theorems for the arising optimization minimax and maximin problems.

[†]Deceased.

¹Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, ul. S. Kovalevskoi 16, Yekaterinburg, 620990 Russia; Institute of Mathematics and Computer Science, Ural Federal University, pr. Lenina 51, Yekaterinburg, 620000 Russia
email: popld@imm.uran.ru

1. PROBLEM STATEMENT AND BASIC ASSUMPTIONS

Consider the linear programming problem

$$\max \{(c, x) \mid Ax \leq b, x \geq 0\} \quad (1.1)$$

and the problem dual to it

$$\min \{(b, y) \mid A^T y \geq c, y \geq 0\}. \quad (1.2)$$

Here, the $m \times n$ numerical matrix A and vectors b and c (of the corresponding dimension) are given, and the vectors x and y of primal and dual variables, respectively, are to be found. Brackets denote the scalar product of vectors. For brevity, we will denote the columns of the matrix A by A_1, A_2, \dots, A_n and its rows by a_1, a_2, \dots, a_m .

Let the original problems be feasible. Denote their common optimal value by $\bar{\gamma}$ and their optimal sets by \bar{X} and \bar{Y} , respectively. It is well known that any pairs of optimal vectors $\bar{x} \in \bar{X}$ and $\bar{y} \in \bar{Y}$ (and only they) form saddle points of the Lagrange function of the original problem

$$F(x, y) = (c, x) - (Ax - b, y) \quad (= (b, y) - (A^T y - c, x))$$

with respect to the domain $\Omega = \{[x, y] \mid x \geq 0, y \geq 0\}$, i.e., $\bar{x} \geq 0$ and $\bar{y} \geq 0$, and the following two-sided inequalities hold:

$$F(x, \bar{y}) \leq F(\bar{x}, \bar{y}) \leq F(\bar{x}, y) \quad \forall x \geq 0 \quad \forall y \geq 0. \quad (1.3)$$

The mentioned equivalence of problems (1.1), (1.2) to saddle point search problem (1.3) is a productive platform for a wide range of linear programming computational methods. However, in almost all such methods (see the references), the Lagrange function is modified in order to impart to it certain useful computational properties. Usually, either regularizing or penalty terms are included in the Lagrange function; in the case of a penalty, functions of an exterior penalty are mainly used. Below, we include penalty terms of logarithmic barrier type in the original saddle function. This allows us, first, to strengthen the convexity-concavity properties of the classical Lagrange function and, second, to constructively take into account the nonnegativity requirement for the variables. As a result, we obtain a problem of searching for a saddle point with respect to the whole space of its variables. We also strengthen the assumptions concerning the original problems and require the existence of Slater's primal and dual points x_S and y_S , respectively. These are points satisfying the conditions $Ax_S < b$, $x_S > 0$ and $A^T y_S > c$, $y_S > 0$. Due to this assumption, the optimal sets \bar{X} and \bar{Y} of the original problem and of the problem dual to it are not only nonempty but also bounded. The above condition is usually imposed in investigations on central path methods [9].

2. APPLICATION OF BARRIERS IN A SADDLE POINT PROBLEM

Let $\epsilon > 0$ be a small numerical parameter. Let us construct the extended Lagrange function of problems (1.1), (1.2)

$$F_\epsilon(x, y) = (c, x) - (Ax - b, y) + \epsilon \sum_{i=1}^n \ln x_i - \epsilon \sum_{j=1}^m \ln y_j \quad (2.1)$$

and consider the problem of searching for its saddle point $[\bar{x}_\epsilon, \bar{y}_\epsilon] > 0$ defined as a (positive) solution of the infinite system of inequalities

$$F_\epsilon(x, \bar{y}_\epsilon) \leq F_\epsilon(\bar{x}_\epsilon, \bar{y}_\epsilon) \leq F_\epsilon(\bar{x}_\epsilon, y) \quad \forall x > 0 \quad \forall y > 0.$$

Let us emphasize that the requirement of the positivity of the primal and dual variables is taken into account here by means of the inclusion of barrier type terms in the Lagrange function. As shown below, the required saddle points lie inside the admissible domain of function (2.1), which makes the problem under consideration similar to saddle point search problems for smooth convex-concave² functions with respect to the whole space.

Let us discuss existence conditions for saddle points of function (2.1) and analyze their connection with the solution of original problems (1.1), (1.2).

Let us introduce the function

$$\Phi_\epsilon(x) = \inf_{y>0} F_\epsilon(x, y).$$

It is easy to see that, for all positive x ,

$$\Phi_\epsilon(x) = \begin{cases} (c, x) + \epsilon \sum_{j=1}^m \ln(b_j - (a_j, x)) + \epsilon \sum_{i=1}^n \ln x_i - m(\ln \epsilon - 1)\epsilon & \text{if } Ax < b, \\ -\infty & \text{otherwise.} \end{cases} \quad (2.2)$$

Indeed, if $(a_{j_0}, x) - b_{j_0} \geq 0$ for at least for one index j_0 , then, due to the unbounded growth of y_{j_0} , the value $F_\epsilon(x, y)$ can be made arbitrarily small, which explains the lower line in (2.2). In the case when $(a_j, x) - b_j < 0$ for all j , the equation

$$\nabla_y F_\epsilon(x, y) = b - Ax - \epsilon \begin{pmatrix} 1/y_1 \\ 1/y_2 \\ \dots \\ 1/y_m \end{pmatrix} = 0$$

is feasible with respect to positive values of the dual variables. The solution $y^\epsilon(x)$ has components

$$y_j^\epsilon(x) = \epsilon / (b_j - (a_j, x)) > 0 \quad (j = 1, \dots, m) \quad (2.3)$$

and defines a minimum point of the function $F_\epsilon(x, \cdot)$ in the dual variables. To validate the first line in (2.2), we should substitute the components of this point into the expression for $F_\epsilon(x, y)$.

Thus, the function $\Phi_\epsilon(x)$ is defined and finite inside the admissible domain of primal problem (1.1) and differs from the classical barrier function associated with the primal problem only by the constant term $m(\ln \epsilon - 1)\epsilon$. The general theory of logarithmic barrier functions [9] yields the following statement.

Assertion 1. *Let the constraints of problems (1.1), (1.2) satisfy Slater's condition. Then, for any $\epsilon > 0$, there exists a unique maximum point $\bar{x}_\epsilon > 0$ of the function $\Phi_\epsilon(x)$, and this point is the (unique) solution of the system of nonlinear equations*

$$\nabla \Phi_\epsilon(x) = c - \epsilon \sum_{j=1}^m a_j (b_j - (a_j, x))^{-1} + \epsilon \begin{pmatrix} 1/x_1 \\ 1/x_2 \\ \dots \\ 1/x_n \end{pmatrix} = 0. \quad (2.4)$$

²The function $F_\epsilon(x, y)$ is convex in dual variables and concave in primal variables.

The points \bar{x}_ϵ for $\epsilon \in (0, +\infty)$ form the so-called *central path* of problem (1.1).

We apply a similar argument to the function

$$\Psi_\epsilon(y) = \sup_{x>0} F_\epsilon(x, y).$$

It is easy to see that, for all positive y

$$\Psi_\epsilon(y) = \begin{cases} (b, y) - \epsilon \sum_{i=1}^n \ln((A_i, y) - c_i) - \epsilon \sum_{j=1}^m \ln y_j + n(\ln \epsilon - 1)\epsilon & \text{if } A^T y > c, \\ +\infty & \text{otherwise.} \end{cases} \quad (2.5)$$

Indeed, if $(A_{i_0}, y) - c_{i_0} \leq 0$ for at least one value index i_0 , then due to the unbounded growth of x_{i_0} , the value $F_\epsilon(x, y)$ can be made arbitrarily large, which explains the lower line in (2.5). In the case when $(A_i, y) - c_i > 0$ for all i , the equation

$$\nabla_x F_\epsilon(x, y) = c - A^T y + \epsilon \begin{pmatrix} 1/x_1 \\ 1/x_2 \\ \dots \\ 1/x_n \end{pmatrix} = 0$$

is feasible with respect to positive values of the primal variables. This solution $x^\epsilon(y)$ has components

$$x_i^\epsilon(y) = \epsilon / ((A_i, x) - c_i) > 0 \quad (i = 1, \dots, n) \quad (2.6)$$

and defines a maximum point of the function $F_\epsilon(\cdot, y)$ in the primal variables. To validate the upper line of (2.5), we should substitute the components of this point into the expression for $F_\epsilon(x, y)$.

Thus, the function $\Psi_\epsilon(y)$ is defined and finite inside the admissible domain of dual problem (1.2) and differs from the classical barrier function associated with the dual problem only the constant term $n(\ln \epsilon - 1)\epsilon$. The general theory of logarithmic barrier functions yields the following statement.

Assertion 2. *Let the constraints of problems (1.1), (1.2) satisfy Slater's condition. Then, for any $\epsilon > 0$, there exists a unique minimum point $\bar{y}_\epsilon > 0$ of the function $\Psi_\epsilon(y)$, and this point is the (unique) solution of the system of nonlinear equations*

$$\nabla \Psi_\epsilon(y) = b - \epsilon \sum_{i=1}^n A_i ((A_i, y) - c_i)^{-1} - \epsilon \begin{pmatrix} 1/y_1 \\ 1/y_2 \\ \dots \\ 1/y_m \end{pmatrix} = 0. \quad (2.7)$$

The points \bar{y}_ϵ for $\epsilon \in (0, +\infty)$ form the so-called *central path* of problem (1.2).

Now, let \bar{x}_ϵ and \bar{y}_ϵ be taken from Assertions 1 and 2. The pair $[\bar{x}_\epsilon, \bar{y}_\epsilon] > 0$ is in fact a saddle point of function (2.1). To show this, note first that

$$\max_{x>0} \Phi_\epsilon(x) = \Phi_\epsilon(\bar{x}_\epsilon) = F_\epsilon(\bar{x}_\epsilon, y^\epsilon(\bar{x}_\epsilon)),$$

where (see (2.3))

$$y^\epsilon(\bar{x}_\epsilon)_j = \epsilon / (b_j - (a_j, \bar{x}_\epsilon)) > 0 \quad (j = 1, \dots, m). \quad (2.8)$$

However, it is easy to see that the components of the vector $y^\epsilon(\bar{x}_\epsilon)$ satisfy system of nonlinear equations (2.7). Indeed, in view of (2.4) and (2.8), we have

$$\nabla \Phi_\epsilon(x) \Big|_{x=\bar{x}_\epsilon} = c - \underbrace{\epsilon \sum_{j=1}^m a_j (b_j - (a_j, \bar{x}_\epsilon))^{-1}}_{c - A^T y^\epsilon(\bar{x}_\epsilon)} + \epsilon \begin{pmatrix} 1/\bar{x}_1^\epsilon \\ 1/\bar{x}_2^\epsilon \\ \dots \\ 1/\bar{x}_n^\epsilon \end{pmatrix} = 0,$$

which yields componentwise

$$(A_i, y^\epsilon(\bar{x}_\epsilon)) - c_i = \epsilon / \bar{x}_\epsilon^i > 0 \quad (i = 1, \dots, n).$$

This implies the required equality (we use (2.8) once more)

$$\nabla \Psi_\epsilon(y) \Big|_{y=y^\epsilon(\bar{x}_\epsilon)} = b - \underbrace{\epsilon \sum_{i=1}^n A_i ((A_i, y^\epsilon(\bar{x}_\epsilon)) - c_i)^{-1}}_{b - A\bar{x}_\epsilon} \epsilon \begin{pmatrix} 1/y_1^\epsilon(\bar{x}_\epsilon) \\ 1/y_2^\epsilon(\bar{x}_\epsilon) \\ \dots \\ 1/y_m^\epsilon(\bar{x}_\epsilon) \end{pmatrix} = b - A\bar{x}_\epsilon - b + A\bar{x}_\epsilon = 0.$$

Consequently, $y^\epsilon(\bar{x}_\epsilon) = \bar{y}_\epsilon$ and

$$\max_{x>0} \Phi_\epsilon(x) = \Phi_\epsilon(\bar{x}_\epsilon) = F_\epsilon(\bar{x}_\epsilon, y^\epsilon(\bar{x}_\epsilon)) = F_\epsilon(\bar{x}_\epsilon, \bar{y}_\epsilon).$$

Similarly,

$$\min_{y>0} \Psi_\epsilon(y) = \Psi_\epsilon(\bar{y}_\epsilon) = F_\epsilon(x^\epsilon(\bar{y}_\epsilon), \bar{y}_\epsilon),$$

where (see (2.6))

$$x^\epsilon(\bar{y}_\epsilon)_i = \epsilon / ((A_i, \bar{y}_\epsilon) - c_i) > 0 \quad (i = 1, \dots, n). \quad (2.9)$$

However, it is easy to see that the components of the vector $x^\epsilon(\bar{y}_\epsilon)$ satisfy system of nonlinear equations (2.4). Indeed, in view of (2.7) and (2.9), we have

$$\nabla \Psi_\epsilon(y) \Big|_{y=y^\epsilon(\bar{x}_\epsilon)} = b - \underbrace{\epsilon \sum_{i=1}^n A_i ((A_i, \bar{y}_\epsilon) - c_i)^{-1}}_{b - A x^\epsilon(\bar{y}_\epsilon)} \epsilon \begin{pmatrix} 1/\bar{y}_1^\epsilon \\ 1/\bar{y}_2^\epsilon \\ \dots \\ 1/\bar{y}_m^\epsilon \end{pmatrix} = 0,$$

which yields componentwise

$$b_j - (a_j, x^\epsilon(\bar{y}_\epsilon)) = \epsilon / \bar{y}_j^\epsilon \quad (j = 1, \dots, m).$$

This implies the required equality (we use (2.9) once more)

$$\nabla \Phi_\epsilon(x) \Big|_{x=x^\epsilon(\bar{y}_\epsilon)} = c - \underbrace{\epsilon \sum_{j=1}^m a_j (b_j - (a_j, x^\epsilon(\bar{y}_\epsilon)))^{-1}}_{c - A^T \bar{y}_\epsilon} - \epsilon \begin{pmatrix} 1/x_1^\epsilon(\bar{y}_\epsilon) \\ 1/x_2^\epsilon(\bar{y}_\epsilon) \\ \dots \\ 1/x_n^\epsilon(\bar{y}_\epsilon) \end{pmatrix} = c - A^T \bar{y}_\epsilon - c + A^T \bar{y}_\epsilon = 0.$$

Consequently, $x^\epsilon(\bar{y}_\epsilon) = \bar{x}_\epsilon$ and

$$\min_{y>0} \Psi_\epsilon(y) = \Psi_\epsilon(\bar{y}_\epsilon) = F_\epsilon(x^\epsilon(\bar{y}_\epsilon), \bar{y}_\epsilon) = F_\epsilon(\bar{x}_\epsilon, \bar{y}_\epsilon).$$

Finally, we have

$$\max_{x>0} \Phi_\epsilon(x) = \min_{y>0} \Psi_\epsilon(y) = F_\epsilon(\bar{x}_\epsilon, \bar{y}_\epsilon).$$

As a result, we can formulate the following statement.

Assertion 3. *Let the constraints of problems (1.1), (1.2) satisfy Slater's condition. Then, for any $\epsilon > 0$, there exists a unique saddle point $[\bar{x}_\epsilon, \bar{y}_\epsilon] > 0$ of the function $F_\epsilon(x, y)$. Moreover, there exist positive constants K_1, K_2 , and K_3 such that*

- (1) $\rho(\bar{x}_\epsilon, \bar{X}) < \epsilon K_1$,
- (2) $\rho(\bar{y}_\epsilon, \bar{Y}) < \epsilon K_2$,
- (3) $|(c, \bar{x}_\epsilon) - (b, \bar{y}_\epsilon)| < \epsilon K_3$.

Here, $\rho(\cdot, \cdot)$ is the Euclidean distance from a point to a set.

Proof. The proof follows from known facts of the theory of logarithmic barrier functions and the central path method.

3. BARRIER FUNCTIONS AND DUALITY

Let us now consider the maximin and minimax search problems for the extended Lagrange function. These problems have the form

$$\bar{\gamma}_\epsilon = \sup_{x>0} \Phi_\epsilon(x) = \sup_{x>0} \inf_{y>0} F_\epsilon(x, y)$$

and

$$\bar{\gamma}_\epsilon = \inf_{y>0} \Psi_\epsilon(y) = \inf_{y>0} \sup_{x>0} F_\epsilon(x, y).$$

In view of the results of the previous sections, these problems can be represented as convex programming problems with equality constraints:

$$\max \left\{ (c, x) + \epsilon \sum_{j=1}^m \ln u_j + \epsilon \sum_{i=1}^n \ln x_i - m(\ln \epsilon - 1)\epsilon \mid Ax + u = b \right\} \quad (3.1)$$

and

$$\min \left\{ (b, y) - \epsilon \sum_{i=1}^n \ln v_i + \epsilon \sum_{j=1}^m \ln y_j + n(\ln \epsilon - 1)\epsilon \mid A^T y - v = c \right\}. \quad (3.2)$$

Assertion 4. *Convex programming problems (3.1) and (3.2) are dual to each other in the classical sense and are in the relation of perfect duality.*

Indeed, write the usual Lagrange function of problem (3.1)

$$L(x, u, w) = (c, x) + \epsilon \sum_{j=1}^m \ln u_j + \epsilon \sum_{i=1}^n \ln x_i - m(\ln \epsilon - 1)\epsilon - (w, Ax + u - b).$$

Evidently, problem (3.1) can be considered as its maximin

$$\bar{\gamma}_\epsilon = \sup_{x>0, u>0} \inf_w L(x, u, w).$$

Let us consider the classical dual (minimax) problem

$$\inf_w \sup_{x>0, u>0} L(x, u, w) \quad (3.3)$$

and investigate the properties of its target function

$$\zeta(w) = \sup_{x>0, u>0} L(x, u, w).$$

It is easy to see that, under the assumption $A^T w - c > 0$ and $w > 0$, the maximum points of $L(x, u, w)$ in the primal variables can be found from the equations

$$\nabla_x L(x, u, w) = c + \epsilon \begin{pmatrix} 1/x_1 \\ 1/x_2 \\ \dots \\ 1/x_n \end{pmatrix} - A^T w = 0, \quad \nabla_u L(x, u, w) = \epsilon \begin{pmatrix} 1/u_1 \\ 1/u_2 \\ \dots \\ 1/u_m \end{pmatrix} - w = 0,$$

which gives vectors $x(w)$ and $u(w)$ with the components

$$x_i(w) = \epsilon / ((A_i, w) - c_i) > 0 \quad (i = 1, \dots, n), \quad u_j(w) = \epsilon / w_j > 0 \quad (j = 1, \dots, m)$$

and the value of the maximum

$$\begin{aligned} \zeta(w) &= L(x(w), u(w), w) \\ &= (b, w) + (c - A^T w, x(w)) - (w, u(w)) + \epsilon \sum_{j=1}^m \ln u_j(w) + \epsilon \sum_{i=1}^n \ln x_i(w) - m(\ln \epsilon - 1)\epsilon \\ &= (b, w) - \epsilon \sum_{i=1}^n \ln ((A_i, w) - c_i) + \epsilon \sum_{j=1}^m \ln w_j + n(\ln \epsilon - 1)\epsilon = \Psi_\epsilon(w). \end{aligned}$$

For all other w , obviously, $\zeta(w) = +\infty$. Thus, problem (3.3) (up to notation for the variables) coincides with the problem

$$\bar{\gamma}_\epsilon = \inf_{y>0} \Psi_\epsilon(y)$$

and, thus, with problem (3.2).

A similar argument can be made for problem (3.2), which is the usual dual problem for (3.1). The fact that these problem are in the relation of perfect duality (the reachability and coincidence of their optimal values) was actually shown in the first two sections of this paper.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 10-01-00273) and by the Presidium of the Ural Branch of the Russian Academy of Sciences (project nos. 12-P-1-1016, 12-S-1-1017/1, 12-P-1-1023, and 12-P-1-1034).

REFERENCES

1. I. I. Eremin, *Theory of Linear Optimization* (VSP, Utrecht, 2002).
2. I. I. Eremin and N. N. Astaf'ev, *Introduction to the Theory of Linear and Convex Programming* (Nauka, Moscow, 1976) [in Russian].
3. W. I. Zangwill, *Nonlinear Programming: A Unified Approach* (Prentice-Hall, Englewood Cliffs, NJ, 1969; Sovetsk. Radio, Moscow, 1973).
4. F. P. Vasil'ev, *Numerical Methods for Solving Extremal Problems* (Nauka, Moscow, 1988) [in Russian].
5. Yu. G. Evtushenko, *Methods for Solving Extremal Problems and Their Application in Optimization Systems* (Nauka, Moscow, 1982) [in Russian].
6. K.-H. Elster, R. Reinhardt, M. Schauble, and G. Donath, *Einführung in die Nichtlineare Optimierung* (Teubner, Leipzig, 1977; Nauka, Moscow, 1985).
7. A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968; Mir, Moscow, 1972).
8. V. D. Skarin, "Barrier function method and correction algorithms for improper convex programming problems," Proc. Steklov Inst. Math., Suppl. 2, S120–S134 (2008).
9. C. Roos, T. Terlaky, and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization* (Wiley, Chichester, 1997).

Translated by E. Vasil'eva